Spatio-temporal analysis of dengue fever in Makassar Indonesia a comparison of models basen on CARBayes

by Aswi Aswi

Submission date: 17-Jun-2021 10:33AM (UTC+0700)
Submission ID: 1607817350
File name: Makassar_Indonesia_a_comparison_of_models_basen_on_CARBayes.pdf (823.26K)
Word count: 4349
Character count: 21509

Chapter 9
Spatio-Temporal Analysis of Dengue Fever in Makassar Indonesia: A Comparison of Models Based on CARBayes

Aswi Aswi, Susanna Cramb, Wenbiao Hu, Gentry White, and Kerrie L. Mengersen

Abstract **Background:** Dengue fever is one of the world's most important vector-borne diseases and it is still a major public health problem in the Asia-Pacific region including Indonesia. Makassar is one of the major cities in Indonesia where the incidence of dengue fever is still quite high. Since dengue cases vary between areas and over time, these spatial and temporal components should be taken into consideration. However, unlike many other spatio-temporal contexts, Makassar is comprised of only a small number of areas and data are available over a relatively short timeframe. The aim of this paper is to better understand the spatial and temporal patterns of dengue incidence in Makassar, Indonesia by comparing the performance of six existing spatio-temporal models, taking into account these specific data characteristics (small number of areas and limited small number of time periods) and to select the best model for Makassar dengue dataset.

Methods: Six different Bayesian spatio-temporal conditional autoregressive (ST CAR) models were compared in the context of a substantive case study, namely annual dengue fever incidence in 14 geographic areas of Makassar, Indonesia, during 2002–2015. The candidate models included linear, ANOVA, separate spatial, autoregressive (AR), adaptive and localised approaches. The models were

Aswi · G. White · K. L. Mengersen (✉)
ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, Brisbane, QLD, Australia
e-mail: k.mengersen@qut.edu.au

Cramb
ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, Brisbane, QLD, Australia

Cancer Council Queensland, Brisbane, QLD, Australia

W. Hu
School of Public Health and Social Work, Queensland University of Technology, Brisbane, QLD, Australia

K. L. Mengersen et al. (eds.), *Case Studies in Applied Bayesian Data Science*, Lecture Notes in Mathematics 2259, https://doi.org/10.1007/978-3-030-42553-1_9

implemented using CARBayesST and the goodness of fit was compared using the Deviance Information Criterion (DIC) and Watanabe-Akaike Information Criterion (WAIC).

Results: The six models performed differently in the context of this case study. Among the six models, the spatio-temporal conditional autoregressive localised model had a much better fit than other options in terms of DIC, while the conditional autoregressive model with separate spatial and temporal components performed worst. However, the spatio-temporal CAR AR had a much better fit than other models in terms of WAIC. The different performance of the models may have been influenced by the small number of areas.

Conclusion: Different spatio-temporal models appeared to have a large impact on results. Careful selection of a range of spatio-temporal models is important for assessing the spatial and temporal patterns of dengue fever, especially in a context characterised by relatively few spatial areas and limited time periods.

Keywords Bayesian · Conditional autoregressive priors · CARBayesST · Spatio-temporal models

9.1 Introduction

Despite concerted efforts worldwide, dengue fever remains a serious health problem in the Asia-Pacific region including Indonesia. Makassar, the gateway to eastern Indonesia, is one of the major cities in Indonesia where the incidence of dengue fever is still quite high. However, there is substantial variation in incidence between districts and over time. Although there is strong interest in developing statistical models to estimate and predict dengue incidence, such models need to take these spatial and temporal components into account. Consideration of only the spatial component of disease can identify regions with low or high risk, but not capture anything about temporal variation of risk which is equally crucial. Similarly, focusing only on temporal variation and ignoring important spatial patterns is inadequate for effective understanding or management of the disease.

Some modelling approaches for dengue fever have been conducted in Makassar. However, these models focus on analysing the genomes of dengue viruses using phylogenetic analysis [1], predicting dengue cases using multiple regression [2] and predicting Dengue Haemorrhagic Fever (DHF) epidemics using two different models, a HR2008 model and a persistence model [3]. Spatio-temporal modelling approaches, and in particular Bayesian models, have not been explored yet for Makassar.

A variety of Bayesian spatial and spatio-temporal approaches have used in modelling dengue fever in other locations. A literature search revealed 31 journal articles about Bayesian spatial and spatio-temporal approaches to modelling dengue fever published from January 2000 to November 2017. Most studies adopted a Bayesian model with a spatially structured random effect using an intrinsic CAR prior structure to investigate the relationship between the risk of dengue and selected

covariates [4]. Among the selected studies, only two studies used a generalised linear mixed model (GLMM) with spatial, temporal and spatio-temporal effects [5, 6]. An interesting feature of these studies is the wide disparity in the number of areas used to partition the region of interest; this ranged from 10 to 1490, with less than a quarter (eight studies) focusing on a small number of areas (<30). The number of periods also varied between studies, ranging from 3 months to 32 years, with 18 studies focusing on small time periods of less than 7 years.

There are specific limitations and challenges in the Makassar data, which are common to many datasets. The first challenge is that the spatial data are only available at the district level, and there are only 14 such areas in Makassar. The second challenge is that this dataset has a small number of time periods. This motivates an investigation of available spatio-temporal models that perform well in a context characterised by a small number of areas and small number of time periods. This paper provides a comparison of six existing spatio-temporal models, with the overall aim of better understanding spatial and temporal patterns of dengue incidence in Makassar, Indonesia. It is anticipated that the results of this evaluation will also inform other studies that are similarly characterised by a small number of areas and time periods.

Another consideration in choosing a statistical model for use in this case study is that the model should be easily implemented with publicly available software. This will enhance the potential for the approach to be adopted by public health agencies in Indonesia. As above, such a requirement is not unique to this case study and will have resonance with analysts and agencies in other developing countries.

9.2 Methods

9.2.1 Study Site

Makassar covers an area of 175.77 km^2 divided into 14 districts. A total of 6882 new cases of dengue were registered from 2002 to 2015, with substantial variation between districts and over time; for example, the number of cases rose sharply by 100 percent from 2012 (86 cases) to 2013 (265 cases) in a population of 1.49 million (2012) and 1.51 million (2013) [7].

9.2.2 Models

Six Bayesian spatio-temporal models with different formulations of conditional autoregressive (CAR) priors, namely linear [8, 9], ANOVA [10], separate spatial [11], AR [12], adaptive [13], and localised [14] models, were compared. These models were chosen because they fulfil the modelling requirements described above, in that they include both spatial and temporal components and they are publicly

available in the CARBayesST package [9] in the statistical software package R [15]. All six models are formulated as follows,

$$y_{ij} \sim \text{Poisson}\left(e_{ij}\theta_{ij}\right)$$

$$\log\left(\theta_{ij}\right) = \psi_{ij}$$

where y_{ij} is the observed number of dengue cases in the ith district and jth time period, $i = 1, \ldots, I$; $j = 1, \ldots, J$; e_{ij} and θ_{ij} are, respectively, the expected number of dengue cases in area i time j and the relative risk of dengue (the underlying disease rate); and ψ_{ij} is a latent component for area i and time j involving one or more sets of spatio-temporally autocorrelated random effects. Details of each model are given below and also summarised in Table 9.1.

Models were compared using two goodness-of-fit measures, namely Deviance Information Criterion (DIC) [16] and Watanabe-Akaike Information Criterion (WAIC) [17], as well as by comparing the obtained estimates and their precision for each area.

9.2.2.1 Spatio Temporal CAR Linear Model

This model is suitable for estimating which areas have increasing or decreasing linear trends in the response over time. Here,

$$\psi_{ij} = \alpha_1 + u_i + (\beta + \delta_i)\frac{j - \overline{j}}{J}$$

where $\boldsymbol{u}$ and $\boldsymbol{\delta}$ denote normally distributed random effects that respectively describe spatial variation and the interaction between spatial and temporal effects. Thus each area i is allowed to have its own linear temporal trend, with a spatially varying intercept and slope $\alpha_1 + u_i$ and $(\beta + \delta_i)$ respectively. The random effects are assigned Leroux priors as follows:

$$(u_i|u_{-i}, \mathbf{W}) \sim \text{N}\left(\frac{\rho_{\text{int}}\sum_{k=1}^{I}\omega_{ik}u_k}{\rho_{\text{int}}\sum_{k=1}^{I}\omega_{ik} + 1 - \rho_{\text{int}}}, \frac{\tau^2_{\text{int}}}{\rho_{\text{int}}\sum_{k=1}^{I}\omega_{ik} + 1 - \rho_{\text{int}}}\right),$$

$$(\delta_i|\delta_{-i}, \mathbf{W}) \sim \text{N}\left(\frac{\rho_{\text{slo}}\sum_{k=1}^{I}\omega_{ik}\delta_k}{\rho_{\text{slo}}\sum_{k=1}^{I}\omega_{ik} + 1 - \rho_{\text{slo}}}, \frac{\tau^2_{\text{slo}}}{\rho_{\text{slo}}\sum_{k=1}^{I}\omega_{ik} + 1 - \rho_{\text{slo}}}\right).$$

Here, the elements of the adjacency matrix $\boldsymbol{W} = (\omega_{ik})$ represent the closeness between areas i and k, such that

$$\omega_{ik} = 1 \text{ if } i, k \text{ are adjacent}, \omega_{ik} = 0 \text{ otherwise};$$

Table 9.1 Summary of the spatio-temporal structure for the six models

Models	Structure ψ_{ij}	Prior			Additional information
		Spatial	Temporal	Spatio-temporal	
ST CAR linear	Intercept +spatial effect (for all time) +temporal effect (for all areas) +space-time interaction	CAR Leroux [18]	Linear	CAR Leroux	
ST CAR ANOVA	Spatial effect (for all time) +temporal effect (for all areas) + independent space-time interaction	CAR Leroux	CAR Leroux	Independence	
ST CAR separate spatial	Separate Spatial effect (for each time) +temporal effect (for all areas)	CAR Leroux	CAR Leroux	–	
ST CAR AR	Spatial effect (for each time)			CAR Leroux + AR (1)	-Temporal autocorrelation is induced via the mean $\rho_T u_{j-1}$, Spatial autocorrelation is induced by the variance $\tau^2\mathbf{Q}(\mathbf{W}, \rho_S)^{-1}$ -Single level ρ_S using CAR Leroux
ST CAR adaptive	Spatial effect (for each time)			CAR Leroux + AR (1)	-Model structure is the same as CAR AR, but this allows for localised spatial autocorrelation ρ_S using CAR Leroux
ST CAR localised	Spatial effect (for each time) + cluster component λ			ICAR +AR (1)	Model structure is the same as CAR AR except a component λ and random effects $\mathbf{u}$ are modelled with $\rho_S = 1$ (ICAR)

AR(1) first order autoregressive, ICAR intrinsic conditional autoregressive, ST spatio-temporal

α_1 is the mean log incidence over all areas; β is the mean linear temporal trend over all areas; τ^2_{int} and τ^2_{slo} are precision terms associated respectively with the intercept and slope of the regression; and ρ_{int}, ρ_{slo} are parameters of spatial dependence with values in the interval [0,1]. The Bayesian model is completed by specifying priors for the hyperparameters; in this study, the default priors in the CARBayes package were evaluated and deemed to be suitable, i.e., τ^2_{int}, $\tau^2_{\text{slo}} \sim$ Inverse-Gamma (1, 0.01); ρ_{int}, $\rho_{\text{slo}}\sim$ Uniform (0, 1); $\beta\sim$N(0, 1000).

9.2.2.2 Spatio Temporal CAR ANOVA Model

This model is suitable for estimating overall temporal trends and spatial patterns. Here,

$$\psi_{ij} = u_i + \delta_j + \gamma_{ij}$$

where $\boldsymbol{u}$ denotes the spatial random effects over all time periods; $\boldsymbol{\delta}$ denotes the temporal random effect over all spatial units and $\boldsymbol{\gamma}$ denotes the space-time random interaction. The priors for the first two of these terms are as follows:

$$(u_i|u_{-i}, \mathbf{W}) \sim \text{N}\left(\frac{\rho_S \sum_{k=1}^{I} \omega_{ik} u_k}{\rho_S \sum_{k=1}^{I} \omega_{ik} + 1 - \rho_S}, \frac{\tau_S^2}{\rho_S \sum_{k=1}^{I} \omega_{ik} + 1 - \rho_S}\right),$$

$$(\delta_j|\delta_{-j}, \mathbf{D}) \sim \text{N}\left(\frac{\rho_T \sum_{k=1}^{J} d_{jk} \delta_k}{\rho_T \sum_{k=1}^{J} d_{jk} + 1 - \rho_T}, \frac{\tau_T^2}{\rho_T \sum_{k=1}^{J} d_{jk} + 1 - \rho_T}\right),$$

where the adjacency matrix $\mathbf{D} = (d_{jk})$ represents the closeness between times j and k, where k is the time immediately before or after j so that $d_{jk} = 1$ if $|k\text{-}j| = 1$ and $d_{jk} = 0$ otherwise. An independent normal prior is assigned to the last term,

$$\gamma_{ij} \sim \text{N}\left(0, \tau_\gamma^2\right).$$

As above, default priors were used for the remaining parameters, so that

$$\tau_S^2, \tau_T^2, \tau_\gamma^2 \sim \text{Inverse-Gamma}\,(1, 0.01),$$

$$\rho_S, \rho_T \sim \text{Uniform}\,(0, 1).$$

9.2.2.3 Spatio Temporal CAR Separate Spatial Model

This model is suitable for estimating overall temporal trends and to what extent the spatial variation has changed over time. Here,

$$\psi_{ij} = u_{ij} + \delta_j,$$

$$\left(u_{ij}|u_{-ij}, \mathbf{W}\right) \sim \mathrm{N}\left(\frac{\rho_S \sum_{k=1}^{I} \omega_{ik} u_{kj}}{\rho_S \sum_{k=1}^{I} \omega_{ik} + 1 - \rho_S}, \frac{\tau_j^2}{\rho_S \sum_{k=1}^{I} \omega_{ik} + 1 - \rho_S}\right),$$

$$\left(\delta_j|\delta_{-j}, \mathbf{D}\right) \sim \mathrm{N}\left(\frac{\rho_T \sum_{k=1}^{J} d_{jk} \delta_k}{\rho_T \sum_{k=1}^{J} d_{jk} + 1 - \rho_T}, \frac{\tau_T^2}{\rho_T \sum_{k=1}^{J} d_{jk} + 1 - \rho_T}\right),$$

$u_j = (u_{1j}, u_{2j}, \ldots u_{Ij})$ are separate spatial effects at each time period j.
$\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_J)$ are temporal random effects over all the spatial areas i.

$$\tau_1^2, \ldots, \tau_J^2, \tau_T^2 \sim \text{Inverse-Gamma}\,(1, 0.01)$$

$$\rho_S, \rho_T \sim \text{Uniform}\,(0, 1)\,.$$

9.2.2.4 Spatio Temporal CAR AR Model

This model is suitable for estimating the evolution of the spatial response surface over time without forcing it to be the same for each time. It has a single level of spatial dependence controlled by ρ_S, so that

$$\psi_{ij} = u_{ij},$$

$$\left(u_j|u_{j-1},\right) \sim \mathrm{N}\left(\rho_T u_{j-1}, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_S)^{-1}\right) \qquad j = 2, \ldots . J,$$

$$u_1 \sim \mathrm{N}\left(\mathbf{0}, \tau^2 \mathbf{Q}(\mathbf{W}, \rho_S)^{-1}\right)$$

$$\tau^2 \sim \text{Inverse-Gamma}\,(1, 0.01)$$

$$\rho_S, \rho_T \sim \text{Uniform}\,(0, 1)\,.$$

9.2.2.5 Spatio Temporal CAR Adaptive Model

This model is an extension of spatio temporal CAR AR to allow for spatially adaptive smoothing (localised spatial autocorrelation), noting that ST CAR AR has only a single level of spatial dependence. This model is suitable when the residual spatial autocorrelation in the response is consistent over time but has

a localised structure. The model structure is the same as CAR AR but nonzero (spatial) elements of neighbourhood matrix (W) can vary locally.

9.2.2.6 Spatio Temporal CAR Localised Model

This model is suitable for identifying clusters of areas that exhibit elevated values of the response compared with their geographical and temporal neighbours. This model structure is the same as CAR AR, but there is an additional cluster component λ and random effects $\mathbf{u}$ are modelled with $\rho_S = 1$ (ICAR). This model is similar to ST adaptive, in that both avoid the restrictive assumption that two areas that are close together must have similar estimates. The differences between the ST CAR localised and ST CAR adaptive models is that ST CAR localised captures any step-changes in the response via the mean function, but ST CAR adaptive captures any step changes via the correlation structure (via **W**). Here,

$$\psi_{ij} = u_{ij} + \lambda_{Z_{ij}},$$

$$\left(u_j | u_{j-1},\right) \sim \mathrm{N}\left(\rho_T u_{j-1},\ \tau^2 \mathbf{Q}(\mathbf{W})^{-1}\right) \qquad j = 2, \ldots . J,$$

$$u_1 \sim \mathrm{N}\left(\mathbf{0}, \tau^2 \mathbf{Q}(\mathbf{W})^{-1}\right)$$

$$\tau^2 \sim \text{Inverse-Gamma}\,(1, 0.01)$$

$$\rho_T \sim \text{Uniform}\,(0, 1)\,.$$

$\lambda_{Z_{ij}}$ is a piecewise constant clustering or intercept component,
$\lambda_k \sim$ Uniform $(\lambda_{k\text{-}1}, \lambda_{k+1})$ for $k = 1, 2, \ldots, G$

$$f\left(Z_{ij} | Z_{i,j-1}\right) = \frac{\exp\left(-\delta\left[\left(Z_{ij} - Z_{i,j-1}\right)^2 + \left(Z_{ij} - G^*\right)^2\right]\right)}{\sum_{r=1}^{G} \exp\left(-\delta\left[\left(r - Z_{i,j-1}\right)^2 + \left(r - G^*\right)^2\right]\right)}$$

for $j = 2, \ldots, J$

$$f\left(Z_{i1}\right) = \frac{\exp\left(-\delta\left(Z_{i1} - G^*\right)^2\right)}{\sum_{r=1}^{G} \exp\left(-\delta(r - G^*)^2\right)}$$

$\delta \sim$ Uniform (1, 10) where δ is the penalty parameter.

9.2.3 Case Study

Annual dengue fever incidence data for Makassar, Indonesia (14 geographic areas) during 2002–2015 were obtained from the Health Office of Makassar, South Sulawesi Province. An ethics exemption to use these datasets was obtained from QUT (exemption number: 1700000479) as it involves the use of existing collections of data that contain only non-identifiable data about human beings.

9.3 Results

9.3.1 Dengue Data

The descriptive analysis and the plot of the number of Makassar dengue cases from 2002 to 2015 can be seen in Table 9.2 and Fig. 9.1 respectively.

All models were fit using the CARBayesST package version 2.5.1 [9] in R version 3.3.3 or [15]. Posterior estimates and inferences were based on 100,000 MCMC samples collected after a burn in of 20,000 samples.

Crude risk estimates of dengue based on a raw SIR (Standardized incidence ratio) model were calculated for each area and represent the risk of being diagnosed with

Table 9.2 Descriptive analysis of dengue cases from 2002 to 2015

Year	Min	1st Qu	Median	Mean	3rd Qu	Max	Var
2002	14.00	44.00	86.00	104.80	98.00	419.00	10622.80
2003	11.00	40.75	61.00	82.43	73.75	251.00	5233.19
2004	10.00	20.50	33.50	45.50	56.25	178.00	1929.96
2005	11.00	26.00	49.50	63.71	71.00	236.00	3747.60
2006	19.00	27.75	55.50	60.86	65.75	209.00	2278.90
2007	4.00	15.25	27.50	32.64	47.25	80.00	550.09
2008	2.00	11.50	14.00	18.93	19.75	59.00	250.53
2009	3.00	9.00	15.00	18.29	20.00	68.00	286.06
2010	1.00	6.25	8.50	13.21	16.75	45.00	147.26
2011	1.00	2.00	3.00	6.07	9.25	16.00	29.46
2012	0.00	2.00	6.00	6.14	8.75	17.00	24.44
2013	4.00	9.00	13.50	18.93	27.75	52.00	210.22
2014	0.00	3.25	6.50	9.93	13.75	41.00	112.38
2015	2.00	4.75	8.00	10.14	14.00	26.00	50.59

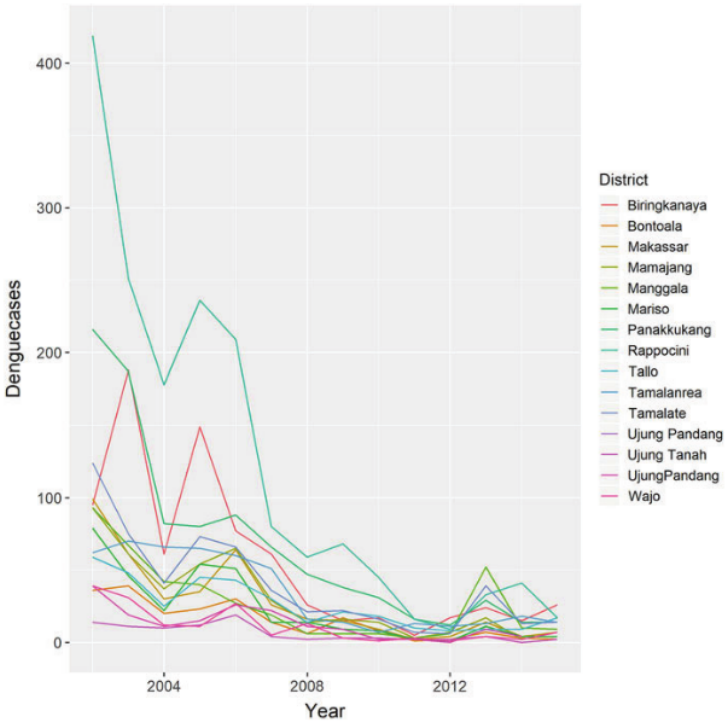

Fig. 9.1 The number of dengue cases in Makassar from 2002 to 2015

dengue fever and are depicted in Fig. 9.2. SIR is the ratio of the observed number of disease cases (y_{ij}) to the expected number of cases (e_{ij}) [19].

$$SIR_{ij} = \frac{y_{ij}}{e_{ij}}$$

It is apparent that there is substantial variation in dengue incidence between districts and over time.

The spatio-temporal CAR localised model with G = 2 had substantially better model fit as demonstrated by the smallest DIC (Table 9.3) followed by the ST CAR AR and ST CAR adaptive models. In contrast, the spatio-temporal CAR separate

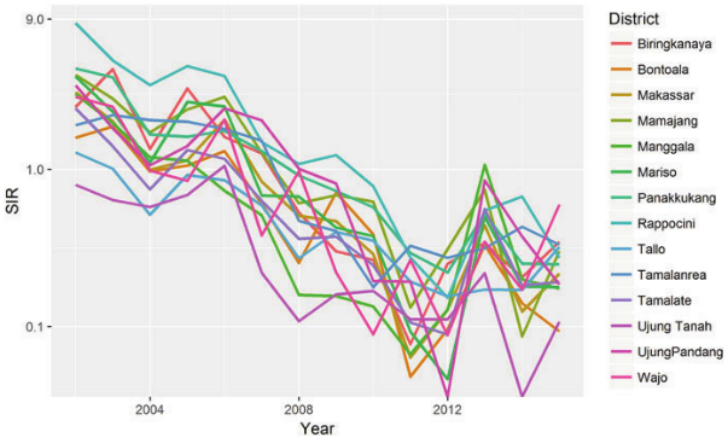

Fig. 9.2 Crude risk estimates (raw SIR model)

Table 9.3 DIC and WAIC for six models using dengue fever data for Makassar

Model	DIC	WAIC	Time (seconds)
ST CAR linear	2012.91	2230.25	44.40
ST CAR ANOVA	9374.02	45374.63	54.70
ST CAR separate spatial	9499.17	Inf	71.80
ST CAR AR	1632.36	**1884.21**	**33.00**
ST CAR Adaptive	1923.21	2111.74	162.90
ST CAR localised, G = 2	**1367.07**	1927.39	117.10
ST CAR localised, G = 3	1438.99	1892.07	128.80

The smallest DIC, WAIC of models and time to run the models are shown in bold

spatial model had the largest DIC. These two models (ST CAR localised, and ST CAR separate spatial) had very different estimates in certain regions and time periods (Fig. 9.3).

Under the preferred spatio temporal localised model with G = 2 (meaning a maximum of two clusters are allowed), most years had two clusters, but a few years had only one cluster (Fig. 9.4 and Table 9.4). Figure 9.5 shows that the overall Standardised Incidence ratios (SIR) have been decreasing over time for all areas but, as discussed above, there was a lot of variation/fluctuation from year to year.

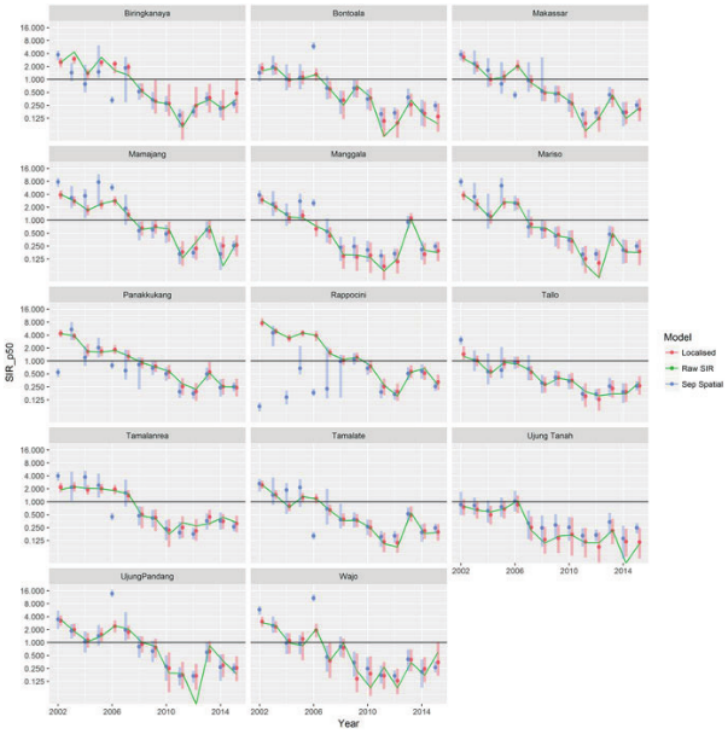

Fig. 9.3 SIR plot under the ST CAR localised model (corresponding to the smallest DIC), ST separate spatial model (corresponding to the largest DIC) for every area, with associated 95% credible intervals, and raw SIR

9.4 Discussion

Six different Bayesian spatio-temporal conditional autoregressive (ST CAR) models were compared by applying them to dengue incidence data from Makassar, Indonesia. The different structures, similarities and dissimilarities of the models have been summarized. The ST CAR localised model with G = 2 proposed by Lee and Lawson [14] performed the best based on the DIC goodness of fit measure, followed by the ST CAR AR and ST adaptive models. However, the ST CAR AR model proposed by Rushworth et al. [12] performed best in terms of WAIC and the computing time required. This is reasonable as the spatio-temporal random effect structure of the ST CAR localised model is the same as the ST CAR AR model except for an additional

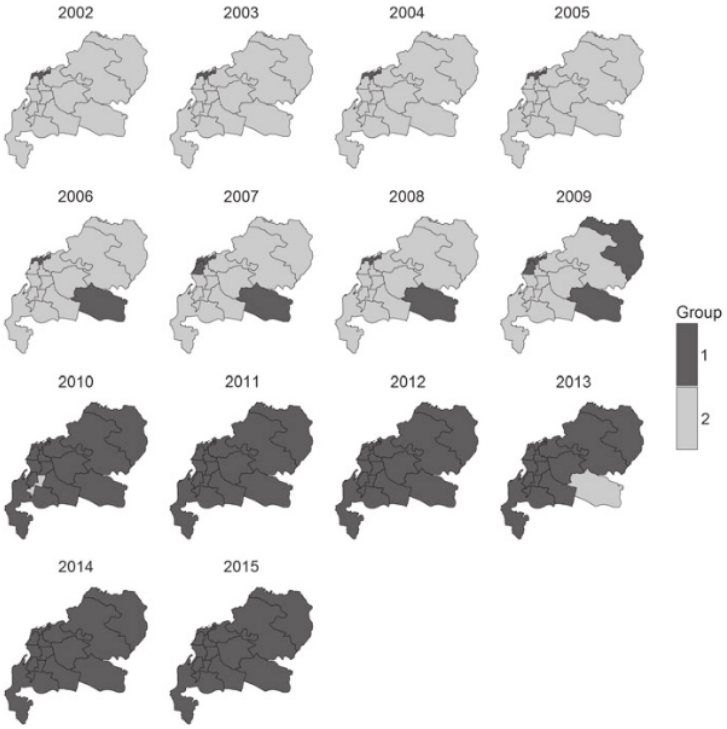

Fig. 9.4 Localised maps obtained under the spatio-temporal localised model with G = 2 local areas

Table 9.4 Districts included in each group under the spatio-temporal localised model

Year	Group 1	Group 2
2002	Ujung Tanah	All districts, except Ujung Tanah
2003	Ujung Tanah	All districts, except Ujung Tanah
2004	Ujung Tanah	All districts, except Ujung Tanah
2005	Ujung Tanah	All districts, except Ujung Tanah
2006	Manggala, Ujung Tanah	All districts, except Manggala, Ujung Tanah
2007	Manggala, Wajo, Ujung Tanah	All districts, except Manggala, Ujung Tanah
2008	Manggala, Ujung Tanah	All districts, except Manggala, Ujung Tanah
2009	Manggala, Ujung Tanah, Wajo, Biringkanaya	All districts, except Manggala, Ujung Tanah, Wajo and Biringkanaya
2010	All districts except Mamajang	Mamajang
2011	All districts	–
2012	All districts	–
2013	All districts, except Manggala	Manggala
2014	All districts	–
2015	All districts	–

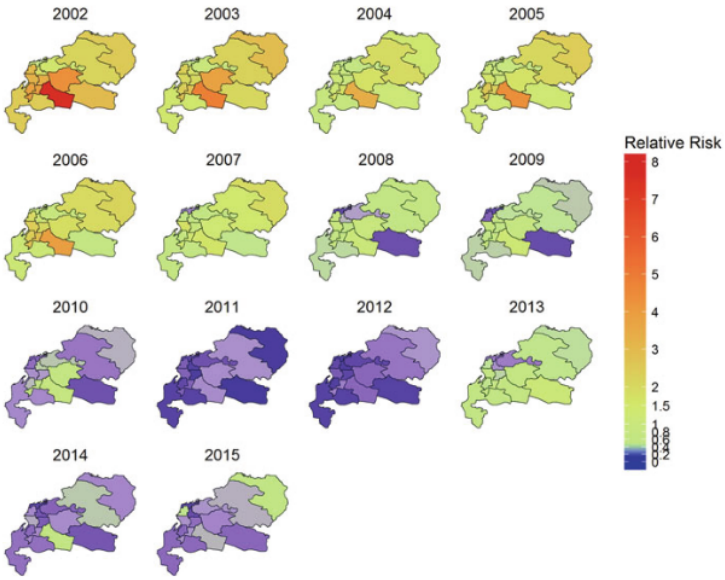

Fig. 9.5 Relative Risk maps obtained under the spatio-temporal localised model with G = 2

cluster component. In contrast, the ST CAR separate spatial model proposed by Napier et al. [11] performed the worst in terms of both DIC and WAIC. This may have been influenced by the small number of areas and time periods.

9.5 Conclusion

Bayesian CAR models can allow for different representations of spatial, temporal and spatio-temporal patterns. Results from the case study showed that the choice of model can have a large impact on goodness of fit, and that a spatio-temporal CAR localised model with G = 2 spatial groups provided the best fit in terms of DIC. Careful exploration of a range of models is important, especially when there are few areas and few time periods. The study motivates future research to provide more general insight into the behaviour of Bayesian spatio-temporal CAR models when the disease rate and degrees of spatio-temporal autocorrelation varies over different numbers of areas and time periods.

Acknowledgments This work was supported by the ARC Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS) and Queensland University of Technology (QUT).

The authors acknowledge the City Health Department of Makassar, South Sulawesi Province for providing and permitting analysis of the dengue fever dataset used in the case study presented in this paper.

References

1. R.T. Sasmono, I. Wahid, H. Trimarsanto, B. Yohan, S. Wahyuni, M. Hertanto, et al., Genomic analysis and growth characteristic of dengue viruses from Makassar, Indonesia. Infect. Genet. Evol. **32**, 165–177 (2015)
2. H. Halide, P. Ridd, A predictive model for dengue hemorrhagic fever epidemics. Int. J. Environ. Health Res. **18**(4), 253–265 (2008)
3. H. Halide, Household protection against dengue hemorrhagic fever epidemics in coastal city of Makassar. J. Coast. Dev. **13**(3), 195–204 (2011)
4. A. Aswi, S.M. Cramb, P. Moraga, K. Mengersen, Bayesian spatial and spatio-temporal approaches to modelling dengue fever: a systematic review. Epidemiol. Infect. 1–14 (2018). https://doi.org/10.1017/S0950268818002807 Epub 10/29
5. S.P. Wijayanti, T. Porphyre, M. Chase-Topping, S.M. Rainey, M. McFarlane, E. Schnettler, et al., The Importance of socio-economic versus environmental risk factors for reported dengue cases in Java, Indonesia. PLoS Negl. Trop. Dis. **10**(9), e0004964 (2016). https://doi.org/10.1371/journal.pntd.0004964
6. D. Martínez-Bello, A. López-Quílez, A. Prieto, Spatiotemporal modeling of relative risk of dengue disease in Colombia. Stoch. Env. Res. Risk A. **32**(6), 1587–1601 (2018). https://doi.org/10.1007/s00477-017-1461-5
7. D. Kesehatan, *Propil Kesehatan Kota Makassar* (Dinas Kesehatan Makassar, 2012)
8. L. Bernardinelli, D. Clayton, C. Pascutto, C. Montomoli, M. Ghislandi, M. Songini, Bayesian analysis of space—time variation in disease risk. Stat. Med. **14**(21–22), 2433–2443 (1995)

9. D. Lee, A. Rushworth, G. Napier, Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package. J. Stat. Soft. **84**(9), 1–39 (2018)
10. L. Knorr-Held, Bayesian modelling of inseparable space-time variation in disease risk. Stat. Med. **19**(17–18), 2555–2567 (1999). https://doi.org/10.1002/1097-0258
11. G. Napier, D. Lee, C. Robertson, A. Lawson, K. Pollock, A model to estimate the impact of changes in MMR vaccine uptake on inequalities in measles susceptibility in Scotland. Stat. Methods Med. Res. **25**(4), 1185–1200 (2016). https://doi.org/10.1177/0962280216660420
12. A. Rushworth, D. Lee, R. Mitchell, A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. Spat. Spatio-Temporal Epidemiol. **10**(C), 29–38 (2014). https://doi.org/10.1016/j.sste.2014.05.001
13. A. Rushworth, D. Lee, C. Sarran, An adaptive spatio-temporal smoothing model for estimating trends and step changes in disease risk. J. R. Stat. Soc. Ser. C **66**(1), 141–157 (2017)
14. D. Lee, A. Lawson, Quantifying the spatial inequality and temporal trends in maternal smoking rates in Glasgow. Ann. Appl. Stat. **10**(3), 1427–1446 (2016)
15. Team RC, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria)
16. D.J. Spiegelhalter, N.G. Best, B.P. Carlin, A. Van Der Linde, Bayesian measures of model complexity and fit. J. R. Stat. Soc. Ser. B (Stat. Methodol.) **64**(4), 583–639 (2002)
17. S. Watanabe, Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. **11**, 3571–3594 (2010)
18. B.G. Leroux, X. Lei, N. Breslow, Statistical Models, in *Epidemiology, the Environment, and Clinical Trials*, ed. by M. E. Halloran, D. Berry, (New York, Springer, 2000)
19. A. Lawson, in *Disease Mapping with WinBUGS and MLwiN*, ed. by W. J. Browne, C. L. Vidal Rodeiro, (Wiley, Hoboken, NJ, 2003), p. 277

Spatio-temporal analysis of dengue fever in Makassar Indonesia a comparison of models basen on CARBayes

ORIGINALITY REPORT

15%	15%	2%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.cambridge.org Internet Source	4%
2	Lee, Duncan, Rushworth, Alastair, Napier, Gary. "Spatio-temporal areal unit modeling in R with conditional autoregressive priors using the CARBayesST package", 'Foundation for Open Access Statistic', 2018 Internet Source	3%
3	infoscience.epfl.ch Internet Source	2%
4	Eprints.qut.edu.au Internet Source	2%
5	ij-healthgeographics.biomedcentral.com Internet Source	2%
6	theses.gla.ac.uk Internet Source	2%

Exclude quotes	On	Exclude matches	< 2%
Exclude bibliography	On		